Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia - Social and Behavioral Sciences 56 (2012) 594 – 601
Procedia
Social and Behavioral Sciences

International Conference on Teaching and Learning in Higher Education (ICTLHE 2012) in conjunction with RCEE & RHED 2012

An Analysis of the Relationship between Effective Teaching and Effective Learning at UTP

Mariyamni Awang[a,*], Balbir Singh[b], Iskandar Dzulkarnain[c]

[a,c]*Petroleum Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Perak, Malaysia.*
[b]*Fundamental and Applied Science Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Perak, Malaysia.*

Abstract

Literature abounds with research on effective teaching and student assessments. Findings are often similar and almost as many are contradictory. One way of defining effective teaching is using Student Evaluation of Teaching (SET) which surveys students' satisfaction on the lecturer's performance. In this paper, we analyse data from the Universiti Teknologi PETRONAS (UTP), taken from 53 courses ranging from foundation up to final year courses and covering humanities to core engineering courses. We investigate the relationship between the lecturers' ratings which is assumed to represent teaching effectiveness and the average class grade of the courses which is assumed to represent learning effectiveness. Our analysis finds a correlation of 39% overall, with minor variations for the variables studied. An obvious conclusion is that SET alone is not adequate for assessing teaching effectiveness and other aspects of teaching and learning are needed.

Keywords: student evaluation; class size; course discipline; course level

1. Introduction

Publications on effectiveness of teaching and SET are voluminous. The proponents of SET have written on the reliability and validity of SET with statistics to prove their theory (Centra, 1993;Marsh & Roche, 1997). Many others do not completely reject students' surveys; they question the use of precise numerical ratings (Wolfer & Johnson, 2003) or preferring descriptive terms (Abrami, d'Appollonia, & Rosenfield, 1997). Liaw & Goh (2003) report on the negative effect of large classes of Economics courses to lecturers' ratings. Low (1999) finds that students' understandings of the questions are different than the intention of the questions.

* Corresponding author, Tel.: +605-368-7049
E-mail address: mariyamni_awang@petronas.com.my

doi:10.1016/j.sbspro.2012.09.693

At UTP, students' surveys are conducted at mid-semester and end of semester before the mid semester test and final examination. Therefore, students' grades or marks obtained in the mid- semester tests do not affect the first survey and the outcomes are used for improving teaching. The end of semester survey results may be affected by the marks garnered earlier on. The questions in the survey are given in Table 1. A rating of 1 minimum to 7 maximum is given and a rating of less than 5 is considered unsatisfactory, resulting in the lecturer being tasked to attend a series of training programs such as effective teaching classes. At times, the heads of department will conduct lecture observations. The survey results are useful in assisting a lecturer to identify his or her weakness in a general area of teaching a course; however, the scant one line question is not helpful in dealing with specific problems. For example, are the students' sufficiently prepared for the course? Is their English adequate? All of these factors can affect a student's performance and also a lecturer's expectations from the class. SET too may be perceived as a threat to lecturers especially inexperienced ones into being lenient with their grades or allowing students to control the class. Several researchers such as Centra (2005) do not find any significance when lecturers are lenient with the grading.

The SET being the main assessment tool at UTP seems to give undue power to students who more often than not do not have the experience or the knowledge to assess. For example, in question 2 of the survey (Table 1), students are asked if the lecturer knows his/her subject. A student's comprehension will most likely determine his or her response rather than the actual knowledge possessed by the lecturer. Consequently, the value of the question and response for improving teaching is not significant. Students' survey is also one of the two main indicators of effective or good teaching during the yearly performance assessment.

Table 1.Questions used for students' evaluation of lecturers

Question Number	Questions
1	Quality of presentation
2	Knowledge on subject matter
3	Interaction with students
4	Ability to sustain students interest
5	Pace of instruction
6	Value of activities and exercises

We do not know of any research that has been conducted at UTP on the significance of students' evaluation to teaching or learning effectiveness. Considering such importance is attached to the students' survey or SET, we aim to find a relationship between the perceived effectiveness as seen by students and the actual grades that they achieved as an indicator of learning.

2. Universiti Teknologi PETRONAS Academic Programmes

The University is a small private institution that has focused on engineering education since its inception in 1997. However lately, applied science and humanities programmes have been introduced. Students are screened stringently resulting in only students with high grades in science and technology being admitted to the University. Consequently, a bias may exist when students rate the humanities lecturers. The University has an academic year consisting of three terms or semesters. Students are required to enroll in any two semesters per year.

Foundation year is equivalent to matriculation or pre-university level in other institutions. The programmes groups students into the engineering stream and technology stream.

2.1. Courses analysed

Eighteen courses from the Management and Humanities Faculty are analysed. These courses are compulsory and aim to help UTP produce well-rounded graduates. Twenty-one core courses and fourteen foundation level courses are analysed using the average class grade and the number of respondents, although lower than the class size, it is still an indication of the class size.

3. Data and Analyses

In this study, the foundation studies data is obtained from the engineering stream and the undergraduate data is taken from the Petroleum Engineering and Geosciences programmes. Non-technical courses offered by the Management and Humanities Department (M&H) are given in Table 3.

The results are summarized in Tables 2, 3 and 4.

Table 2. Students' responses and grades obtained for subjects taken in Foundation Studies

Semester	Lecturer	Course Name	No. of responses	Rating (7)	Mean Grade	Mode Grade
Jan	Lecturer A	Science 1	85	6.8	C+	C+
Jan	Lecturer B	Science 2	96	6.1	B	B
Jan	Lecturer C	Science 1	64	6.2	C+	C+
Jan	Lecturer E	Science 4	57	6.4	B-	B
Jan	Lecturer F	Science 5	74	6.3	B-	C+
Jan	Lecturer G	Science 5	84	6.7	B-	C+
May	Lecturer D	Science 3	161	5.2	B-	B
Sept	Lecturer A	Science 2	48	6.9	B	B
Sept	Lecturer B	Science 2	68	6.4	B	B
Sept	Lecturer C	Science 2	76	6.3	B	B
Sept	Lecturer D	Science 4	81	6.1	B	B
Sept	Lecturer E	Science 4	75	6.4	B	B
Sept	Lecturer F	Science 5	81	6.1	B-	B
Sept	Lecturer G	Science 5	99	6.5	B-	B

In Table 2, the comparison of Science 1 subject shows that although the lecturers are highly rated, but the average course grades are at C+. In fact Lecturer A is scoring 8.82 % higher than Lecturer C for the Science I course. The number of responses for Lecturer A is 25% more, and yet the score is 8.82 % better. The less number

of students should be better, as smaller class usually benefits more. Effective learning is usually correlated with small class size. Lecturer A and C both taught Science 2 course in the September semester. Lecturer A's rating improved to 6.9 as the number of responses is smaller, as compared to Science I class that was taught in January semester. The average course grade also improved to B. Lecturer C also improved in course rating, by 1.6% although number of responses increased by 15.8%. Lecturer B, who taught Science 2 in both semesters, for different groups of students, got an increase of 3.17% in course rating, as the number of responses reduces by 20.8%. Both Lecturers F and G show consistency in the reduction of course rating as the number of responses increases. Although the rating is lower for the Science 5 subject, but the mode course grade actually is better. The outcomes shown for Lecturer D confirm that there is very small correlation between effective teaching and course rating. Even though the course rating is low at 5.2, effective learning still takes place as evident from the average course grade.

Table 3.Students' responses and grades obtained for subjects taken in Petroleum Engineering and Petroleum Geosciences programmes

Semester	Lecturer	Course Name	No. of responses	Rating (7)	Mean Grade	Mode Grade
Jan	Lecturer K	GS 1	18	5.9	B-	B
Jan	Lecturer K	PE 1	23	4.6	B-	B
Jan	Lecturer S	GS 2	39	5.8	C+	B
Jan	Lecturer T	GS 3	20	6.2	A-	A-
Jan	Lecturer T	GS 4	39	6.3	B+	A-
Jan	Lecturer T	GS 5	25	6.3	B+	B+
Jan	Lecturer U	GS 2	56	4.8	C+	B
Jan	Lecturer U	PE 2	66	5	C+	C+
Jan	Lecturer V	PE 3	96	6.4	B+	A
Jan	Lecturer W	PE 4	27	6.5	B+	B
May	Lecturer H	PE 5	60	6.8	B-	B
May	Lecturer I	GS 6	21	6.4	B	B
May	Lecturer J	GS 7	26	5.8	B-	B
May	Lecturer K	PE 1	50	5.9	B-	B
May	Lecturer L	GS 7	40	5.9	B-	B
May	Lecturer M	GS 8	26	6.3	A-	B+
May	Lecturer N	GS 9	19	5.9	A	A
May	Lecturer O	PE 6	55	4.5	C	C+
May	Lecturer P	PE 6	63	6.2	C	C+
May	Lecturer Q	PE 7	13	6.5	B+	B+
May	Lecturer R	PE 8	70	4.3	D+	C+

Comparing lecturers' ratings for all courses taught and the corresponding grades may illustrate important point about including other sources of data to assess teaching and learning effectiveness. In Table 3, Lecturer K taught PE1 in both semesters. Although May semester registered a marked increase of 54% in class size, the average

grade is consistent at B-. This occurs despite improvement of 28.3% in the rating. One might argue that class size as determined from the number of respondents might be pivotal in determining the rating but this is not the case here. Lecturer K also taught GS1 which is smaller and yet the rating is similar to course PE1 for May semester. Despite the difference in class size and similarity in rating, there is no marked improvement in student learning as seen from the average grades. Clearly this does not show a complete picture of the lecturer's teaching effectiveness.

Similar argument also applies in the case of Lecturer T. Courses GS3, GS4 and GS5 were all offered in January semester and taught by Lecturer T. In all classes, Lecturer T obtained a consistent rating above 6 and average grades above B+. The SET results might be interpreted as another ordinary performance by the lecturer if not for the fact that the teaching methods employed are not entirely classroom-based. The lecturer used combination of outdoor field trips, practical hands-on exercise and other activities to deliver the course content without depending solely on lectures. Admittedly, this generated students' interest in the subjects. Nonetheless, this dynamic interaction is not captured explicitly in SET.

Still a more balanced perspective would be gained by looking at the ratings of low performing lecturers U, O and R. Even though the classes are medium-sized (50-70 students), the ratings obtained are all below 5. These coincide with the average grades which are below C+. This is one instance where SET may be useful as instrument to solicit feedback on the lecturer's teaching effectiveness. However, SET alone would not give a meaningful feedback if it is not complemented by other sources of measurement. In this instance, lecturer's course portfolio, course outcome survey and peer evaluations would be suggested to ascertain the area where improvement is needed.

Table 4. Students' responses and grades obtained for subjects offered in Management and Humanities Department

Semester	Lecturer	Course Name	No. of responses	Rating (7)	Mean Grade	Mode Grade
Jan	Lecturer X	MH1	28	6.9	B+	B
Jan	Lecturer ZD	MH2	15	5.5	A-	A-
Jan	Lecturer ZE	MH3	26	6.9	A-	B+
Jan	Lecturer ZA	MH4	62	5.3	A-	A
Jan	Lecturer ZB	MH3	17	6.7	A-	B+
Jan	Lecturer ZG	MH5	44	6.4	B-	B
Jan	Lecturer ZH	MH6	11	5.4	B+	B
Jan	Lecturer ZI	MH7	108	6.2	B+	B
Jan	Lecturer ZJ	MH3	29	6.9	A-	B+
May	Lecturer Y	MH1	39	6.8	B+	B
May	Lecturer Z	MH8	45	6.7	B-	B
May	Lecturer ZA	MH9	70	5.6	B+	B
May	Lecturer ZA	MH4	50	5.8	B-	B
May	Lecturer ZB	MH3	23	6.4	A-	A-
May	Lecturer ZC	MH10	42	6.3	B+	A
May	Lecturer ZC	MH9	17	5.2	B+	B
May	Lecturer ZE	MH3	45	6.6	A-	A-
May	Lecturer ZF	MH11	28	6.5	A-	A-

The course ratings and their corresponding grades for M&H Department are shown in Table 4. By looking at the course ratings of several lecturers who scored above 6.0, it seems to suggest that smaller class (as indicated

by number of responses) leads to improved ratings and consequently better grades. One example is Lecturer ZB and ZE who each maintains their course ratings above 6.0 in course MH3 despite a slight drop of 4.5% and 4.3% in May. Although the class size in May increased 35.3% and 73% respectively, these can still be considered as small class since the number of students is below 50. In both cases, the average grade remained consistent at A-. The same can be said on Lecturer X and Y who taught MH1 in January and May consecutively. Despite the small class, both scored 6.9 and 6.8 respectively with the average grade maintained at B+. Yet another similar example is Lecturer ZF who taught MH12 in May. For a class of 28 students, the average grade is A- with course rating of 6.5.

Nevertheless, the same data also indicate that effective teaching (as indicated by the higher rating) does not necessarily translate to higher grades. Conversely, lower rating does not inevitably mean dismal learning effectiveness. Take Lecturer ZC, ZD and ZH as example. Although each lecturer obtained course rating of 5.2, 5.5 and 5.4 respectively, the lowest mean grade for all is B+ and the highest is A-, which is still acceptable considering the general student population. Even though we said earlier that the data shows good correlation between class size and both effective teaching and learning, there is one exception in the case of Lecturer ZA. The mean grade in the class MH4 which had 62 students is A- and the mode grade A. The high grades suggested the students learnt effectively although the same cannot be said on teaching effectiveness since the lecturer had scored 5.3 on the course rating. Clearly this is one situation where SET is deficient to provide satisfactory explanation.

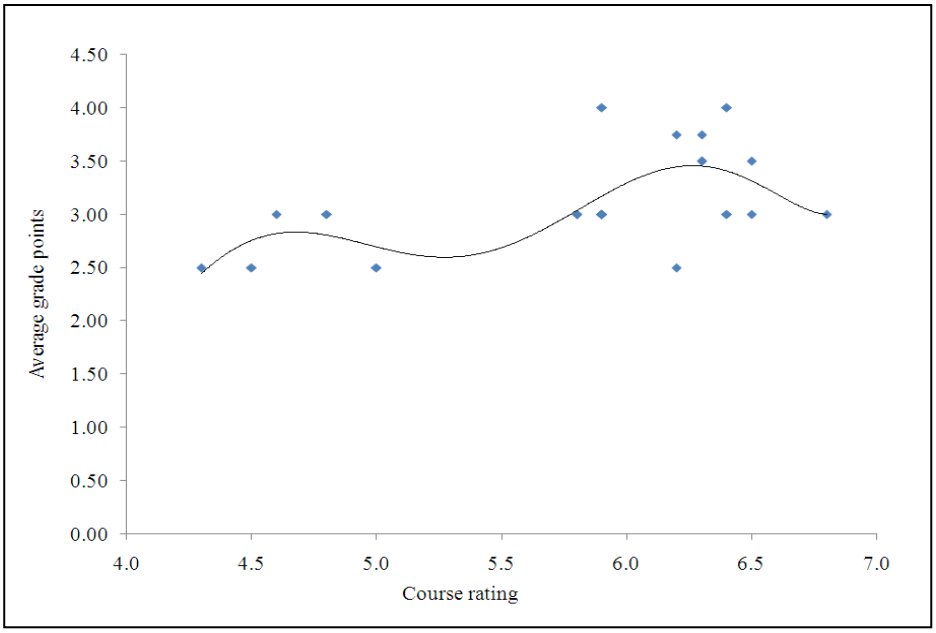

Fig. 1. The average grade points versus course rating for each subject in Petroleum and Geoscience programmes

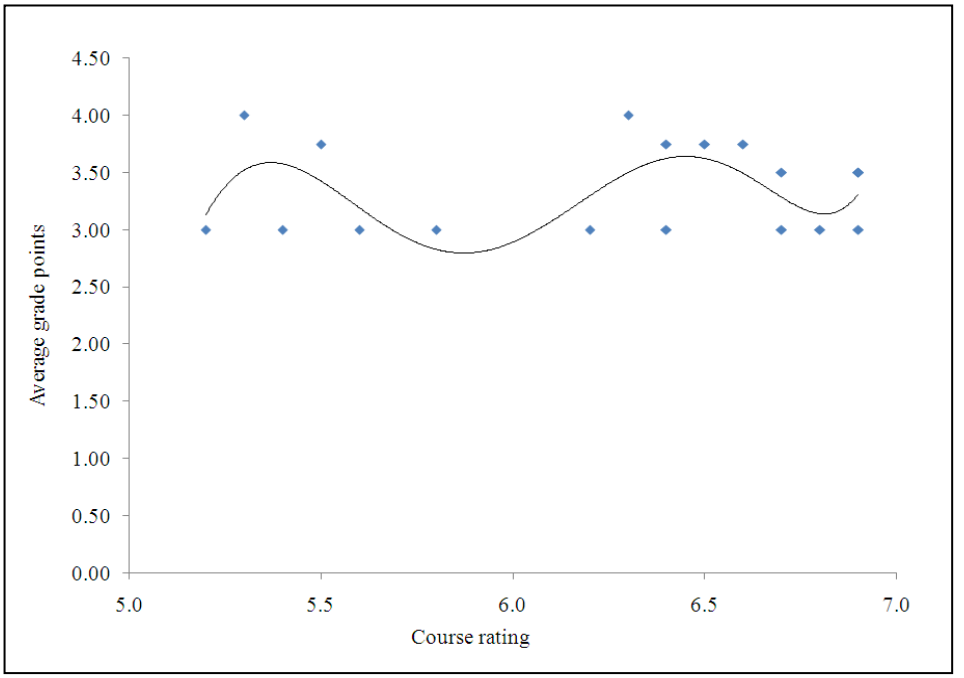

Fig. 2.The average grade points versus course rating for each subject in Management and Humanities Department

In Figure 1 and 2, the grade points are used to analyse whether there is any correlation between the grades and the course ratings. The grade points are applied to the data as given in Table 3 and Table 4. It can be seen from the figures that there is very low correlation between the course ratings and the average grade points. The correlation for the data that is shown in Figure 1 is 39.6% and 28.8% for Figure 2. These are applied for low number of datasets, as required by t-statistics analysis. These results suggest that other factors might influence the students' grades apart from teaching effectiveness. Similarly, learning effectiveness might be due to other reasons beside the lecturer's performance. Nonetheless, these analyses might offer evidence that SET alone cannot be used as the sole criterion for evaluating effective teaching and effective learning.

4. Conclusions

Our analyses show the same results as most researches, that is a correlation of less than 40% is estimated. A lower correlation is obtained for the humanities courses. The class size shows a weak effect. Considering that the student surveys are being used for determining effective teaching and annual performance, a higher correlation should be evident. The authors recommend that a multidimensional approach to effective teaching assessment should be implemented for a fairer evaluation.

Acknowledgements

We wish to thank Universiti Teknologi PETRONAS for enabling us to present this paper and support given by each respective department in preparing this work.

References

Abrami, P. C., d'Appollonia, S., & Rosenfield, S. (1997). The dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry, & J. C. Smart, *Effective teaching in higher education: Research and practice* (pp. 321-367). Bronx, N.Y.: Agathon Press.
Centra, J. A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness.* San Francisco: Jossey-Bass.
Centra, J. A. (2005). *Will Teachers Receive Higher Student Evaluations By Giving Higher Grades and Less Coursework?* Princeton, NJ: Educational Testing Service.
Liaw, S.-H., & Goh, K. L. (2003). Evidence and control of biases in student evaluations of teaching. *International Journal of Educational Management Vol. 17 Iss: 1* , 37-43.
Low, G. (1999). What respondents do with questionnaires: Accounting for incongruity and fluidity. *Applied Linguistics, 20(4)* , 503-533.
Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52 (11)* , 1187-1197.
Wolfer, T. A., & Johnson, M. M. (2003). Re-evaluating student evaluation of teaching: The Teaching Evaluation Form. *Journal of Social Work Education, 39(1)* , 111-121.